Paper No. 14

UNITED STATES PATENT AND TRADEMARK OFFICE
_____

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES
_____

Ex parte RICHARD A. SHARMAN,
and PAUL S. ADAMS
_____

Appeal No. 97-2541
Application 08/369,022[1]
_____

ON BRIEF
_____

Before BARRETT, LEE and RUGGIERO, Administrative Patent Judges.

LEE, Administrative Patent Judge.

DECISION ON APPEAL

This is a decision on appeal under 35 U.S.C. § 134 from the examiner's rejection of claims 1-16. No claim has been allowed.

References relied on by the Examiner

Pate, "Trends in multimedia applications and the network models to support them," Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '90), pp. 317-321, 12/02/90.

_____

[1] Application for patent filed January 5, 1995.

Appeal No. 97-2541
Application 08/369,022

Morse, "Using Computer Conferencing to Improve Productivity in the '90's,' Proceedings of 1988 IEEE Engineering Management Conference, PP. 177-186, 10/24/88.

Shelley et al.   Patent No. 5,345,551   Sep. 6, 1994
 (Shelley)   (Filed Nov.
9,1992)

The Rejections on Appeal

Claims 1-6 and 10-16 stand rejected under 35 U.S.C. § 103

as being unpatentable over Pate and Morse.

Claims 7-9 stand rejected under 35 U.S.C. § 103 as being

unpatentable over Pate, Morse, and Shelley.

The Invention

The invention is directed to a system and method for

conducting a conference between multiple participants each

located at a workstation. Independent claims 1 and 11 are

reproduced below:

     1. A method of textually recording at a workstation
spoken contributions to an audio conference, each participant
in the conference having an associated workstation, the
workstations being linked together by one or more networks,
the method comprising the steps of:

          receiving local speech input at the workstation;

          performing speech recognition on the local
speech input at the workstation to generate a local text
equivalent;

          transmitting the local speech input to the other

2

participant(s) in the conference plus the corresponding text equivalents transmitted from the workstation associated with the respective participant;

storing both the local text equivalents and the text equivalents received from the other workstation(s) in a text file.

11. A system for textually recording at a workstation spoken contributions to an audio conference, each participant in the conference having an associated workstation, the workstations being linked together by one or more networks, the method comprising the steps of:

means for receiving local speech input at the workstation;

means for performing speech recognition on the local speech input at the workstation to generate a local text equivalent;

means for transmitting the local speech input to the other participant(s) in the conference;

means for receiving spoken contributions from the other participant(s) in the conference plus the corresponding text equivalents transmitted from the workstation associated with the respective participant;

means for storing both the local text equivalents and the text equivalents received from the other workstation(s) in a text file.

Opinion

We reverse. Our opinion is based solely on the rationale and position as articulated and advanced by the examiner. A reversal of the rejections on appeal is only an indication of

the lack of merit of the examiner's expressed view and not an

affirmative or blank indication that the claims are

patentable.

Pate generally discusses trends in multimedia

applications. Examples include multimedia mail and multimedia

conferencing. Furthermore, Pate discusses media conversion.

On page 319, in column 2, Pate states:

> Since multimedia applications by design are
> able to handle different types of media, an
> obvious next step is to convert between
> media types. It might be cheaper to store
> or transfer information in a different
> format than the one in which it will be
> used. One example would be speech stored
> and transferred as text that is converted
> back to speech at the destination. As with
> most media conversion, some information is
> lost -- in this case the characteristics of
> the original speaker's voice -- but for
> some applications that is not a problem.
> Media conversion can be especially helpful
> for the disabled. Text can be converted to
> speech for the blind, speech can be
> converted to text or even to video clips of
> simple sign language gestures for the deaf.

The appellants argue that Pate contemplates media

conversion only as alternatives, depending on specific

criteria such as cost reduction in transmission. The

appellants point out that the claimed invention requires the

receipt of both (1) speech data from other audio conference

participants, and (2) the converted text equivalent to that

data, and that Pate nowhere suggests the dual-receipt feature

of the claimed invention.

To satisfy the requirement of receiving both speech and

its text equivalent, the examiner cites to the following

disclosure in Pate on page 318, in column 2, paragraph 2 [1]:

> This application is an extension of the
> text-based electronic mail model to
> incorporate audio, video, and graphics.

It should also be noted that the above-quoted text is

immediately preceded by the caption "Multimedia Mail:". The

examiner further quotes Pate on page 318, in column 2,

paragraph 4:

> NEC in Japan is also investigating
> multimedia conferencing incorporating
> graphics, still images, text, voice, and
> hand drawn figures transmitted using
> satellite communication and ISDN.

The disclosure cited by the examiner, however, is

generic, and would not have reasonably suggested the specific

requirement of receiving both the speech data as well as its

corresponding "text equivalent" as the appellants have

claimed. The general teaching of an extension of the text-

based electronic mail model to incorporate audio, video, and

graphics, would not have reasonably suggested sending both

speech and its "equivalent" text. For instance, the

incorporated audio may simply add explanation to a particular

portion of the text, as is described in Pate on page 318,

column 2, paragraph 1:

> With these features [incorporating audio,
> video, and data] it is possible to create a
> message with a graphics overlay or
> explanatory text of audio synchronized to
> appropriate points in the message.

Explanatory audio material added to a text message is not

the same as speech data accompanied by its equivalent text

arrived at by local speech recognition on the audio speech

input. Note that claim 1 requires performing speech

recognition on the local speech input at the workstation to

generate a local text equivalent, and claim 11 recites a means

for performing speech recognition on the local speech input at

the work station to generate a local text equivalent. Thus,

corresponding text equivalent is generated from audio speech

by speech recognition.

Also, with regard to multimedia conferencing purported
being investigated by NEC, the description is not sufficiently
specific. Mere teaching multimedia conferencing incorporating
graphics, still images, text, voice, and hand drawn figures
transmitted using satellite communication and ISDN does not
reasonably suggest sending both audio speech together with its
locally generated text equivalent to other conferees. A
conclusion of obviousness cannot be based on speculation. The
mere fact that the prior art may be modified in the manner
suggested by the examiner does not make the modification
obvious unless the prior art suggested the desirability of the
modification. In re Fritch, 972 F.2d 1260, 1266 n.14, 23
USPQ2d 1780, 1783-84 n.14 (Fed. Cir. 1992); In re Gordon, 733
F.2d 900, 902, 221 USPQ 1125, 1127 (Fed. Cir. 1984).

With regard to claims 3 and 14, which further require the
transmission of the local text equivalent of the local speech
input to the other workstations in the conference, the
examiner is erroneous in concluding that Morse discloses or
reasonably suggests that feature. In Morse, all inputs are
transmitted to a central system and there is no teaching of
any local conversion of speech into text. Without there being

7

any locally generated text equivalent to an audio speech

input, Morse would not have reasonably suggested sending the

local text equivalent to other workstations. In Morse, inputs

are stored in a central repository called a file and a

conference participant accesses the file at its own

convenience. See Morse at page 178, column 1, paragraphs 2

and 4. Morse would not have reasonably suggested either (1)

receiving both speech input and a corresponding local text

equivalent of that speech from another workstation, as is

required by claims 1 and 11, or (2) transmitting the local

text equivalent of a local speech input to other workstations,

as is required by claims 3 and 14. The examiner has

identified no reasonable suggestion from Morse that the dialog

information generated from each station should contain both

audio speech and corresponding locally generated equivalent

text.

As applied by the examiner, Shelley does not make up for

the deficiencies of Pate and Morse. Accordingly, for the

foregoing reasons, we do not sustain the rejection of claims

1-6 and 10-16 under 35 U.S.C. § 103 as being unpatentable over

Pate and Morse,

and we do not sustain the rejection of claims 7-9 under 35

U.S.C. § 103 as being unpatentable over Pate, Morse, and

Shelley.

Conclusion

The rejection of claims 1-6 and 10-16 under 35 U.S.C. §

103 as being unpatentable over Pate and Morse is **reversed**.

The rejection of claims 7-9 under 35 U.S.C. § 103 as

being unpatentable over Pate and Morse and Shelley is

reversed.

REVERSED

LEE E. BARRETT	)
Administrative Patent Judge	)
	)
	)
	) BOARD OF PATENT
JAMESON LEE	) APPEALS AND
Administrative Patent Judge	) INTERFERENCES
	)
	)
	)
JOSEPH RUGGIERO	)
Administrative Patent Judge	)

Appeal No. 97-2541
Application 08/369,022

Stephen T. Keohane
972/B656
IBM CORPORATION
P.O. Box 12195
Research Triangle Park, NC   27709